

Sequence Listing

Sequence Number : 1

Sequence Length : 2578

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Genomic DNA

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

GCATGCCATT AAAAGATGTA ACATTTTACA CTCCAGACGG TAAGGAGGTT GATGAGAAAG 60
CATGGAATTC CCCAACGCAA ACTGTTATTT TCGTGTTAGA GGGGAGCGTA ATGGATGAGA 120
TTAACATCTA TGGAGAGAGA ATTGCGGATG ATTCATTCTT GATAATTCTT AACGCAAATC 180
CCAATAACGT AAAAGTGAAG TTCCCAAAGG GTAAATGGGA ACTAGTTGTT GGTTCCTTATT 240
TGAGAGAGAT AAAACCAGAA GAAAGAATTG TAGAAGGTGA GAAGGAATTG GAAATTGAGG 300
GAAGAACAGC ATTAGTTTAT AGGAGGACAG AACT ATG ATA ATA GGC ACA TAT AGG 355

Met Ile Ile Gly Thr Tyr Arg

1

5

CTG CAA CTC AAT AAG AAA TTC ACT TTT TAC GAT ATA ATA GAA AAT TTG 403

Leu Gln Leu Asn Lys Lys Phe Thr Phe Tyr Asp Ile Ile Gln Asn Leu

10

15

20

Sub
Paper #12.

Delete
Pages 180-240.

GAT TAT TTT AAA GAA TTA GGA GTA TCA CAC CTA TAT CTA TCT CCA ATA	451
Asp Tyr Phe Lys Glu Leu Gly Val Ser His Leu Tyr Leu Ser Pro Ile	
25 30 35	
CTT AAG GCT AGA CCA GGG AGC ACT CAC GGC TAC GAT GTA GTA GAT CAT	499
Leu Lys Ala Arg Pro Gly Ser Thr His Gly Tyr Asp Val Val Asp His	
40 45 50 55	
AGT GAA ATT AAT GAG GAA TTA GGA GGA GAA GAG GGG TGC TTT AAA CTA	547
Ser Glu Ile Asn Glu Glu Leu Gly Gly Glu Glu Gly Cys Phe Lys Leu	
60 65 70	
GTT AAG GAA GCT AAG AGT AGA GGT TTA GAA ATC ATA CAA GAT ATA GTG	595
Val Lys Glu Ala Lys Ser Arg Gly Leu Glu Ile Ile Gln Asp Ile Val	
75 80 85	
CCA AAT CAC ATG GCG GTA CAT CAT ACT AAT TGG AGA CTT ATG GAT CTG	643
Pro Asn His Met Ala Val His His Thr Asn Trp Arg Leu Met Asp Leu	
90 95 100	
TTA AAG AGT TGG AAG AAT AGT AAA TAC TAT AAC TAT TTT GAT CAC TAC	691
Leu Lys Ser Trp Lys Asn Ser Lys Tyr Tyr Asn Tyr Phe Asp His Tyr	
105 110 115	
GAT GAT GAC AAG ATA ATC CTC CCA ATA CTT GAG GAC GAG TTG GAT ACC	739
Asp Asp Asp Lys Ile Ile Leu Pro Ile Leu Glu Asp Glu Leu Asp Thr	
120 125 130 135	
GTT ATA GAT AAG GGA TTG ATA AAA CTA CAG AAG GAT AAT ATA GAG TAC	787
Val Ile Asp Lys Gly Leu Ile Lys Leu Gln Lys Asp Asn Ile Glu Tyr	
140 145 150	

AGA GGG CTT ATA TTA CCT ATA AAT GAT GAA GGA GTT GAA TTC TTG AAA 835
Arg Gly Leu Ile Leu Pro Ile Asn Asp Glu Gly Val Glu Phe Leu Lys

155

160

165

AGG ATT AAT TGC TTT GAT AAT TCA TGT TTA AAG AAA GAG GAT ATA AAG 883
Arg Ile Asn Cys Phe Asp Asn Ser Cys Leu Lys Lys Glu Asp Ile Lys

170

175

180

AAA TTA CTA TTA ATA CAA TAT TAT CAG CTA ACT TAC TGG AAG AAA GGT 931
Lys Leu Leu Leu Ile Gln Tyr Tyr Gln Leu Thr Tyr Trp Lys Lys Gly

185

190

195

TAT CCA AAC TAT AGG AGA TTT TTC GCA GTA AAT GAT TTG ATA GCT GTT 979
Tyr Pro Asn Tyr Arg Arg Phe Phe Ala Val Asn Asp Leu Ile Ala Val

200

205

210

215

AGG GTA GAA TTG GAT GAA GTA TTT AGA GAG TCC CAT GAG ATA ATT GCT 1027
Arg Val Glu Leu Asp Glu Val Phe Arg Glu Ser His Glu Ile Ile Ala

220

225

230

AAG CTA CCA GTT GAC GGT TTA AGA ATT GAC CAC ATA GAT GGA CTA TAT 1075
Lys Leu Pro Val Asp Gly Leu Arg Ile Asp His Ile Asp Gly Leu Tyr

235

240

245

AAC CCT AAG GAG TAT TTA GAT AAG CTA AGA CAG TTA GTA GGA AAT GAT 1123
Asn Pro Lys Glu Tyr Leu Asp Lys Leu Arg Gln Leu Val Gly Asn Asp

250

255

260

AAG ATA ATA TAC GTA GAG AAG ATA TTG TCA ATC AAC GAG AAA TTA AGA 1171
Lys Ile Ile Tyr Val Glu Lys Ile Leu Ser Ile Asn Glu Lys Leu Arg

265

270

275

GAT GAT TGG AAA GTA GAT GGG ACT ACT GGA TAT GAT TTC TTG AAC TAC	1219
Asp Asp Trp Lys Val Asp Gly Thr Thr Gly Tyr Asp Phe Leu Asn Tyr	
280 285 290 295	
GTT AAT ATG CTA TTA GTA GAT GGA AGT GGT GAG GAG GAG TTA ACT AAG	1267
Val Asn Met Leu Leu Val Asp Gly Ser Gly Glu Glu Glu Leu Thr Lys	
300 305 310	
TTT TAT GAG AAT TTC ATT GGA AGG AAA ATC AAT ATA GAC GAG TTA ATA	1315
Phe Tyr Glu Asn Phe Ile Gly Arg Lys Ile Asn Ile Asp Glu Leu Ile	
315 320 325	
ATA CAA AGT AAA AAA TTA GTT GCA AAT CAG TTA TTT AAA GGT GAC ATT	1363
Ile Gln Ser Lys Lys Leu Val Ala Asn Gln Leu Phe Lys Gly Asp Ile	
330 335 340	
GAA AGA TTA AGC AAG TTA CTG AAC GTT AAT TAC GAT TAT TTA GTA GAT	1411
Glu Arg Leu Ser Lys Leu Leu Asn Val Asn Tyr Asp Tyr Leu Val Asp	
345 350 355	
TTT CTA GCA TGT ATG AAA AAA TAC AGG ACT TAT TTA CCA TAT GAG GAT	1459
Phe Leu Ala Cys Met Lys Lys Tyr Arg Thr Tyr Leu Pro Tyr Glu Asp	
360 365 370 375	
ATT AAC GGA ATA AGG GAA TGC GAT AAG GAG GGA AAG TTA AAA GAT GAA	1507
Ile Asn Gly Ile Arg Glu Cys Asp Lys Glu Gly Lys Leu Lys Asp Glu	
380 385 390	
AAG GGA ATC ATG AGA CTC CAA CAA TAC ATG CCA GCA ATC TTC GCT AAG	1555
Lys Gly Ile Met Arg Leu Gln Gln Tyr Met Pro Ala Ile Phe Ala Lys	
395 400 405	

GGC TAT GAG GAT ACT ACC CTC TTC ATC TAC AAT AGA TTA ATT TCC CTT 1603
Gly Tyr Glu Asp Thr Thr Leu Phe Ile Tyr Asn Arg Leu Ile Ser Leu

410

415

420

AAC GAG GTT GGG AGC GAC CTA AGA AGA TTC AGT TTA AGC ATC AAA GAC 1651
Asn Glu Val Gly Ser Asp Leu Arg Arg Phe Ser Leu Ser Ile Lys Asp

425

430

435

TTT CAT AAC TTT AAC CTA AGC AGA GTA AAT ACC ATA TCA ATG AAC ACT 1699
Phe His Asn Phe Asn Leu Ser Arg Val Asn Thr Ile Ser Met Asn Thr

440

445

450

455

CTT TCC ACT CAT GAT ACT AAA TTC AGT GAA GAC GTT AGA GCT AGA ATA 1747
Leu Ser Thr His Asp Thr Lys Phe Ser Glu Asp Val Arg Ala Arg Ile

460

465

470

TCA GTA CTA TCT GAG ATA CCA AAG GAG TGG GAG GAG AGG GTA ATA TAC 1795
Ser Val Leu Ser Glu Ile Pro Lys Glu Trp Glu Glu Arg Val Ile Tyr

475

480

485

TGG CAT GAT TTG TTA AGC CCA AAT ATT GAT AAA AAC GAT GAG TAT AGA 1843
Trp His Asp Leu Leu Arg Pro Asn Ile Asp Lys Asn Asp Glu Tyr Arg

490

495

500

TTT TAT CAA ACA CTT GTG GGA AGT TAC GAG GGA TTT GAT AAT AAG GAG 1891
Phe Tyr Gln Thr Leu Val Gly Ser Tyr Glu Gly Phe Asp Asn Lys Glu

505

510

515

AGA ATT AAG AAC CAC ATG ATT AAG GTC ATA AGA GAA GCT AAG GTA CAT 1939
Arg Ile Lys Asn His Met Ile Lys Val Ile Arg Glu Ala Lys Val His

520

525

530

535

ACA ACG TGG GAA AAT CCT AAT ATA GAG TAT GAA AAG AAG GTT CTG GGT	1987
Thr Thr Trp Glu Asn Pro Asn Ile Glu Tyr Glu Lys Lys Val Leu Gly	
540 545 550	
TTC ATA GAT GAA GTG TTC GAG AAC AGT AAT TTT AGA AAT GAT TTT GAA	2035
Phe Ile Asp Glu Val Phe Glu Asn Ser Asn Phe Arg Asn Asp Phe Glu	
555 560 565	
AAT TTT GAA AAG AAA ATA GTT TAT TTC GGT TAT ATG AAA TCA TTA ATC	2083
Asn Phe Glu Lys Lys Ile Val Tyr Phe Gly Tyr Met Lys Ser Leu Ile	
570 575 580	
GCA ACG ACA CTT AGG TTC CTT TCG CCC GGT GTA CCA GAT ATT TAT CAA	2131
Ala Thr Thr Leu Arg Phe Leu Ser Pro Gly Val Pro Asp Ile Tyr Gln	
585 590 595	
GGA ACT GAA GTT TGG AGA TTC TTA CTT ACA GAC CCA GAT AAC AGA ATG	2179
Gly Thr Glu Val Trp Arg Phe Leu Leu Thr Asp Pro Asp Asn Arg Met	
600 605 610 615	
CCG GTG GAT TTC AAG AAA CTA AAG GAA TTA TTA AAT AAT TTG ACT GAA	2227
Pro Val Asp Phe Lys Lys Leu Lys Glu Leu Leu Asn Asn Leu Thr Glu	
620 625 630	
AAG AAC TTA GAA CTC TCA GAT CCA AGA GTC AAA ATG TTA TAT GTT AAG	2275
Lys Asn Leu Glu Leu Ser Asp Pro Arg Val Lys Met Leu Tyr Val Lys	
635 640 645	
AAA TTG CTA CAG CTT AGA AGA GAG TAC TCA CTA AAC GAT TAT AAA CCA	2323
Lys Leu Leu Gln Leu Arg Arg Glu Tyr Ser Leu Asn Asp Tyr Lys Pro	
650 655 660	

[illegible]

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Met	Ile	Ile	Gly	Thr	Tyr	Arg	Leu	Gln	Leu	Asn	Lys	Lys	Phe	Thr	Phe
1				5					10					15	
Tyr	Asp	Ile	Ile	Glu	Asn	Leu	Asp	Tyr	Phe	Lys	Glu	Leu	Gly	Val	Ser
			20					25						30	
His	Leu	Tyr	Leu	Ser	Pro	Ile	Leu	Lys	Ala	Arg	Pro	Gly	Ser	Thr	His
			35					40						45	
Gly	Tyr	Asp	Val	Val	Asp	His	Ser	Glu	Ile	Asn	Glu	Glu	Leu	Gly	Gly
			50					55						60	
Glu	Glu	Gly	Cys	Phe	Lys	Leu	Val	Lys	Glu	Ala	Lys	Ser	Arg	Gly	Leu
65					70					75					80
Glu	Ile	Ile	Gln	Asp	Ile	Val	Pro	Asn	His	Met	Ala	Val	His	His	Thr
				85					90					95	
Asn	Trp	Arg	Leu	Met	Asp	Leu	Leu	Lys	Ser	Trp	Lys	Asn	Ser	Lys	Tyr
				100					105					110	
Tyr	Asn	Tyr	Phe	Asp	His	Tyr	Asp	Asp	Asp	Lys	Ile	Ile	Leu	Pro	Ile
				115					120					125	
Leu	Glu	Asp	Glu	Leu	Asp	Thr	Val	Ile	Asp	Lys	Gly	Leu	Ile	Lys	Leu
				130					135					140	
Gln	Lys	Asp	Asn	Ile	Glu	Tyr	Arg	Gly	Leu	Ile	Leu	Pro	Ile	Asn	Asp
145					150					155					160
Glu	Gly	Val	Glu	Phe	Leu	Lys	Arg	Ile	Asn	Cys	Phe	Asp	Asn	Ser	Cys
					165					170					175

Leu Lys Lys Glu Asp Ile Lys Lys Leu Leu Leu Ile Gln Tyr Tyr Gln

180

185

190

Leu Thr Tyr Trp Lys Lys Gly Tyr Pro Asn Tyr Arg Arg Phe Phe Ala

195

200

205

Val Asn Asp Leu Ile Ala Val Arg Val Glu Leu Asp Glu Val Phe Arg

210

215

220

Glu Ser His Glu Ile Ile Ala Lys Leu Pro Val Asp Gly Leu Arg Ile

225

230

235

240

Asp His Ile Asp Gly Leu Tyr Asn Pro Lys Glu Tyr Leu Asp Lys Leu

245

250

255

Arg Gln Leu Val Gly Asn Asp Lys Ile Ile Tyr Val Glu Lys Ile Leu

260

265

270

Ser Ile Asn Glu Lys Leu Arg Asp Asp Trp Lys Val Asp Gly Thr Thr

275

280

285

Gly Tyr Asp Phe Leu Asn Tyr Val Asn Met Leu Leu Val Asp Gly Ser

290

295

300

Gly Glu Glu Glu Leu Thr Lys Phe Tyr Glu Asn Phe Ile Gly Arg Lys

305

310

315

320

Ile Asn Ile Asp Glu Leu Ile Ile Gln Ser Lys Lys Leu Val Ala Asn

325

330

335

Gln Leu Phe Lys Gly Asp Ile Glu Arg Leu Ser Lys Leu Leu Asn Val

340

345

350

Asn Tyr Asp Tyr Leu Val Asp Phe Leu Ala Cys Met Lys Lys Tyr Arg

355

360

365

Thr Tyr Leu Pro Tyr Glu Asp Ile Asn Gly Ile Arg Glu Cys Asp Lys			
370	375	380	
Glu Gly Lys Leu Lys Asp Glu Lys Gly Ile Met Arg Leu Gln Gln Tyr			
385	390	395	400
Met Pro Ala Ile Phe Ala Lys Gly Tyr Glu Asp Thr Thr Leu Phe Ile			
	405	410	415
Tyr Asn Arg Leu Ile Ser Leu Asn Glu Val Gly Ser Asp Leu Arg Arg			
	420	425	430
Phe Ser Leu Ser Ile Lys Asp Phe His Asn Phe Asn Leu Ser Arg Val			
	435	440	445
Asn Thr Ile Ser Met Asn Thr Leu Ser Thr His Asp Thr Lys Phe Ser			
	450	455	460
Glu Asp Val Arg Ala Arg Ile Ser Val Leu Ser Glu Ile Pro Lys Glu			
465	470	475	480
Trp Glu Glu Arg Val Ile Tyr Trp His Asp Leu Leu Arg Pro Asn Ile			
	485	490	495
Asp Lys Asn Asp Glu Tyr Arg Phe Tyr Gln Thr Leu Val Gly Ser Tyr			
	500	505	510
Glu Gly Phe Asp Asn Lys Glu Arg Ile Lys Asn His Met Ile Lys Val			
	515	520	525
Ile Arg Glu Ala Lys Val His Thr Thr Trp Glu Asn Pro Asn Ile Glu			
	530	535	540
Tyr Glu Lys Lys Val Leu Gly Phe Ile Asp Glu Val Phe Glu Asn Ser			
545	550	555	560

Asn Phe Arg Asn Asp Phe Glu Asn Phe Glu Lys Lys Ile Val Tyr Phe

565

570

575

Gly Tyr Met Lys Ser Leu Ile Ala Thr Thr Leu Arg Phe Leu Ser Pro

580

585

590

Gly Val Pro Asp Ile Tyr Gln Gly Thr Glu Val Trp Arg Phe Leu Leu

595

600

605

Thr Asp Pro Asp Asn Arg Met Pro Val Asp Phe Lys Lys Leu Lys Glu

610

615

620

Leu Leu Asn Asn Leu Thr Glu Lys Asn Leu Glu Leu Ser Asp Pro Arg

625

630

635

640

Val Lys Met Leu Tyr Val Lys Lys Leu Leu Gln Leu Arg Arg Glu Tyr

645

650

655

Ser Leu Asn Asp Tyr Lys Pro Leu Pro Phe Gly Phe Gln Arg Gly Lys

660

665

670

Val Ala Val Leu Phe Ser Pro Ile Val Thr Arg Glu Val Lys Glu Lys

675

680

685

Ile Ser Ile Arg Gln Lys Ser Val Asp Trp Ile Arg Asn Glu Glu Ile

690

695

700

Ser Ser Gly Glu Tyr Asn Leu Ser Glu Leu Ile Gly Lys His Lys Val

705

710

715

720

Val Ile Leu Thr Glu Lys Arg Glu

725

Sequence Number : 3

Sequence Length : 3467

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Genomic DNA

Original Source

Organism : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

GCTAATAAAC TGAACAATGA GGACCGAATG AATGAAAATT ATAGCTGGAA TTGTGGAGTA	60
GAAGGAGAAA CTAACGATTC TAATATTCTT TATTGTAGAG AAAAACAAG AAGAAATTTT	120
GTAATAACAT TATTTGTTAG CCAAGGTATA CCAATGATCT TAGGGGGAGA CGAAATAGGA	180
AGAACACAAA AAGGCAACAA TAATGCTTTT TGTCAAGATA ATGAGACAAG TTGGTATGAT	240
TGGAACCTTG ATGAAAATCG TCTAAGGTTT CATGATTTTG TGAGGAGACT TACCAATTTT	300
TATAAAGCTC ATCCGATATT TAGGAGGGCT AGATATTTTC AGGGTAAGAA GTTACACCGT	360
TCCCCATTAA AGGATGTGAC GTGGCTAAAA CCTGACGGCA ATGAAGTTGA TGATTCAGTG	420
TGGAATCTC CAACAAATCA TATTATTTAT ATATTAGAGG GAAGTGCTAT CGATGAAATA	480
AATTATAATG GAGAAAGGAT AGCTGACGAC ACTTTTCTAA TTATTTTGAA TGGAGCAAGT	540
ACTAATCTTA AGATAAAAGT ACCTCATGGA AAATGGGAGT TAGTGTTACA TCCTTATCCA	600
CATGAGCCAT CTAACGATAA AAAGATAATA GAAAACAACA AAGAAGTAGA AATAGATGGA	660
AAGACTGCAC TAATTTACAG GAGGATAGAG TTCCAGTGAT ATCAGCAACC TACAGATTAC	720
AGTTAAATAA GAATTTTAAT TTTGGTGACC TAATCGATAA CCTATGCTAT TTTAAGGATT	780

TAGGAGTTTC CCATCTCTAC CTCTCTCCTG TCTTA ATG GCT TCG CCA GGA AGT AAC 836

Met Ala Ser Pro Gly Ser Asn

1

5

CAT GGG TAC GAT GTA ATA GAT CAT TCA AGG ATA AAC GAT GAA CTT GGA 884

His Gly Tyr Asp Val Ile Asp His Ser Arg Ile Asn Asp Glu Leu Gly

10

15

20

GGA GAG AAA GAA TAC AGG AGA TTA ATA GAG ACA GCT CAT ACT ATT GGA 932

Gly Glu Lys Glu Tyr Arg Arg Leu Ile Glu Thr Ala His Thr Ile Gly

25

30

35

TTA GGT ATT ATA CAG GAC ATA GTA CCA AAT CAC ATG GCT GTA AAT TCT 980

Leu Gly Ile Ile Gln Asp Ile Val Pro Asn His Met Ala Val Asn Ser

40

45

50

55

CTA AAT TGG CGA CTA ATG GAT GTA TTA AAA ATG GGT AAA AAG AGT AAA 1028

Leu Asn Trp Arg Leu Met Asp Val Leu Lys Met Gly Lys Lys Ser Lys

60

65

70

TAT TAT ACG TAC TTT GAC TTT TTC CCA GAA GAT GAT AAG ATA CGA TTA 1076

Tyr Tyr Thr Tyr Phe Asp Phe Phe Pro Glu Asp Asp Lys Ile Arg Leu

75

80

85

CCC ATA TTA GGA GAA GAT TTA GAT ACA GTG ATA AGT AAA GGT TTA TTA 1124

Pro Ile Leu Gly Glu Asp Leu Asp Thr Val Ile Ser Lys Gly Leu Leu

90

95

100

AAG ATA GTA AAA GAT GGA GAT GAA TAT TTC CTA GAA TAT TTC AAA TGG 1172

Lys Ile Val Lys Asp Gly Asp Glu Tyr Phe Leu Glu Tyr Phe Lys Trp

105

110

115

AAA CTT CCT CTA ACA GAG GTT GGA AAT GAT ATA TAC GAC ACT TTA CAA	1220
Lys Leu Pro Leu Thr Glu Val Gly Asn Asp Ile Tyr Asp Thr Leu Gln	
120 125 130 135	
AAA CAG AAT TAT ACC CTA ATG TCT TGG AAA AAT CCT CCT AGC TAT AGA	1268
Lys Gln Asn Tyr Thr Leu Met Ser Trp Lys Asn Pro Pro Ser Tyr Arg	
140 145 150	
CGA TTC TTC GAT GTT AAT ACT TTA ATA GGA GTA AAT GTC GAA AAA GAT	1316
Arg Phe Phe Asp Val Asn Thr Leu Ile Gly Val Asn Val Glu Lys Asp	
155 160 165	
CAC GTA TTT CAA GAG TCC CAT TCA AAG ATC TTA GAT TTA GAT GTT GAT	1364
His Val Phe Gln Glu Ser His Ser Lys Ile Leu Asp Leu Asp Val Asp	
170 175 180	
GGC TAT AGA ATT GAT CAT ATT GAT GGA TTA TAT GAT CCT GAG AAA TAT	1412
Gly Tyr Arg Ile Asp His Ile Asp Gly Leu Tyr Asp Pro Glu Lys Tyr	
185 190 195	
ATT AAT GAC CTG AGG TCA ATA ATT AAA AAT AAA ATA ATT ATT GTA GAA	1460
Ile Asn Asp Leu Arg Ser Ile Ile Lys Asn Lys Ile Ile Ile Val Glu	
200 205 210 215	
AAA ATT CTG GGA TTT CAG GAG GAA TTA AAA TTA AAT TCA GAT GGA ACT	1508
Lys Ile Leu Gly Phe Gln Glu Glu Leu Lys Leu Asn Ser Asp Gly Thr	
220 225 230	
ACA GGA TAT GAC TTC TTA AAT TAC TCC AAC TTA CTG TTT AAT TTT AAT	1556
Thr Gly Tyr Asp Phe Leu Asn Tyr Ser Asn Leu Leu Phe Asn Phe Asn	
235 240 245	

CAA GAG ATA ATG GAC AGT ATA TAT GAG AAT TTC ACA GCG GAG AAA ATA	1604
Gln Glu Ile Met Asp Ser Ile Tyr Glu Asn Phe Thr Ala Glu Lys Ile	
250 255 260	
TCT ATA AGT GAA AGT ATA AAG AAA ATA AAA GCG CAA ATA ATT GAT GAG	1652
Ser Ile Ser Glu Ser Ile Lys Lys Ile Lys Ala Gln Ile Ile Asp Glu	
265 270 275	
CTA TTT AGT TAT GAA GTT AAA AGA TTA GCA TCA CAA CTA GGA ATT AGC	1700
Leu Phe Ser Tyr Glu Val Lys Arg Leu Ala Ser Gln Leu Gly Ile Ser	
280 285 290 295	
TAC GAT ATA TTG AGA GAT TAC CTT TCT TGT ATA GAT GTG TAC AGA ACT	1748
Tyr Asp Ile Leu Arg Asp Tyr Leu Ser Cys Ile Asp Val Tyr Arg Thr	
300 305 310	
TAT GCT AAT CAG ATT GTA AAA GAG TGT GAT AAG ACC AAT GAG ATA GAG	1796
Tyr Ala Asn Gln Ile Val Lys Glu Cys Asp Lys Thr Asn Glu Ile Glu	
315 320 325	
GAA GCA ACC AAA AGA AAT CCA GAG GCT TAT ACT AAA TTA CAA CAA TAT	1844
Glu Ala Thr Lys Arg Asn Pro Glu Ala Tyr Thr Lys Leu Gln Gln Tyr	
330 335 340	
ATG CCA GCA GTA TAC GCT AAA GCT TAT GAA GAT ACT TTC CTC TTT AGA	1892
Met Pro Ala Val Tyr Ala Lys Ala Tyr Glu Asp Thr Phe Leu Phe Arg	
345 350 355	
TAC AAT AGA TTA ATA TCC ATA AAT GAG GTT GGA ACC GAT TTA CGA TAT	1940
Tyr Asn Arg Leu Ile Ser Ile Asn Glu Val Gly Ser Asp Leu Arg Tyr	
360 365 370 375	

TAT AAG ATA TCG CCT GAT CAG TTT CAT GTA TTT AAT CAA AAA CGA AGA	1988
Tyr Lys Ile Ser Pro Asp Gln Phe His Val Phe Asn Gln Lys Arg Arg	
380 385 390	
GGA AAA ATC ACA CTA AAT GCC ACT AGC ACA CAT GAT ACT AAG TTT AGT	2036
Gly Lys Ile Thr Leu Asn Ala Thr Ser Thr His Asp Thr Lys Phe Ser	
395 400 405	
GAA GAT GTA AGG ATG AAA ATA AGT GTA TTA AGT GAA TTT CCT GAA GAA	2084
Glu Asp Val Arg Met Lys Ile Ser Val Leu Ser Glu Phe Pro Glu Glu	
410 415 420	
TGG AAA AAT AAG GTC GAG GAA TGG CAT AGT ATC ATA AAT CCA AAG GTA	2132
Trp Lys Asn Lys Val Glu Glu Trp His Ser Ile Ile Asn Pro Lys Val	
425 430 435	
TCA AGA AAT GAT GAA TAT AGA TAT TAT CAG GTT TTA GTG GGA AGT TTT	2180
Ser Arg Asn Asp Glu Tyr Arg Tyr Tyr Gln Val Leu Val Gly Ser Phe	
440 445 450 455	
TAT GAG GGA TTC TCT AAT GAT TTT AAG GAG AGA ATA AAG CAA CAT ATG	2228
Tyr Glu Gly Phe Ser Asn Asp Phe Lys Glu Arg Ile Lys Gln His Met	
460 465 470	
ATA AAA AGT GTC AGA GAA GCT AAG ATA AAT ACC TCA TGG AGA AAT CAA	2276
Ile Lys Ser Val Arg Glu Ala Lys Ile Asn Thr Ser Trp Arg Asn Gln	
475 480 485	
AAT AAA GAA TAT GAA AAT AGA GTA ATG GAA TTA GTG GAA GAA ACT TTT	2324
Asn Lys Glu Tyr Glu Asn Arg Val Met Glu Leu Val Glu Glu Thr Phe	
490 495 500	

ACC AAT AAG GAT TTC ATT AAA AGT TTC ATG AAA TTT GAA AGT AAG ATA	2372
Thr Asn Lys Asp Phe Ile Lys Ser Phe Met Lys Phe Glu Ser Lys Ile	
505 510 515	
AGA AGG ATA GGG ATG ATT AAG AGC TTA TCC TTG GTC GCA TTA AAA ATT	2420
Arg Arg Ile Gly Met Ile Lys Ser Leu Ser Leu Val Ala Leu Lys Ile	
520 525 530 535	
ATG TCA GCC GGT ATA CCT GAT TTT TAT CAG GGA ACA GAA ATA TGG CGA	2468
Met Ser Ala Gly Ile Pro Asp Phe Tyr Gln Gly Thr Glu Ile Trp Arg	
540 545 550	
TAT TTA CTT ACA GAT CCA GAT AAC AGA GTC CCA GTG GAT TTT AAG AAA	2516
Tyr Leu Leu Thr Asp Pro Asp Asn Arg Val Pro Val Asp Phe Lys Lys	
555 560 565	
TTA CAC GAA ATA TTA GAA AAA TCC AAA AAA TTT GAA AAA AAT ATG TTA	2564
Leu His Glu Ile Leu Glu Lys Ser Lys Lys Phe Glu Lys Asn Met Leu	
570 575 580	
GAG TCT ATG GAC GAT GGA AGA ATT AAG ATG TAT TTA ACA TAT AAG CTT	2612
Glu Ser Met Asp Asp Gly Arg Ile Lys Met Tyr Leu Thr Tyr Lys Leu	
585 590 595	
TTA TCC CTA AGA AAA CAG TTG GCT GAG GAT TTT TTA AAG GGC GAG TAT	2660
Leu Ser Leu Arg Lys Gln Leu Ala Glu Asp Phe Leu Lys Gly Glu Tyr	
600 605 610 615	
AAG GCA TTA GAT CTA GAA GAA GGA CTA TGT GGG TTT ATT AGG TTT AAC	2708
Lys Gly Leu Asp Leu Glu Glu Gly Leu Cys Gly Phe Ile Arg Phe Asn	
620 625 630	

AAA ATT TTG GTA ATA ATA AAA ACC AAG GGA AGT GTT AAT TAC AAA CTG	2756
Lys Ile Leu Val Ile Ile Lys Thr Lys Gly Ser Val Asn Tyr Lys Leu	
635 640 645	
AAA CTT GAA GAG GGA GCA ATT TAC ACA GAT GTA TTG ACA GGA GAA GAA	2804
Lys Leu Glu Glu Gly Ala Ile Tyr Thr Asp Val Leu Thr Gly Glu Glu	
650 655 660	
ATT AAA AAA GAG GTA CAG ATT AAT GAG CTA CCT AGG ATA CTA GTT AGA	2852
Ile Lys Lys Glu Val Gln Ile Asn Glu Leu Pro Arg Ile Leu Val Arg	
665 670 675	
ATG TAAGTTATAA TAATCCGATT TTTATGTGAC AAGATTTACG CTTACGAAAA	2905
Met	
680	
GGACTGTTAA ATCAACTTTT ATGTGAATTA TGAAACGTAA ATTATAAGTT TCCTGAGGAT	2965
AAACATATAT ATCTCTATCT CTCATTGATA TCACATGAGT ATTAGATTAA GGGGAAGTAA	3025
TTCTTACGGA CATTGAGGCT GGTTCACAGT ATACTGTAGA ATATGTAATA GGAAAATAAG	3085
AATAGGAACG GACTTAGTCT ACAAATGCCC TAAATGTGAA AAGAAGTATA ACGCATTCTT	3145
CTGTGAAGCA GATGCTAGGG GATTAAAGAA AAAGTGCCCA TACTGTGGTA CTGAACTTGT	3205
CAGTGCAATT TAAGACTCAA ATAGAAGGTA AAAATATTTT TATACTGAAT AATGAGTTGT	3265
TTTACGCTGA TACGGATATA GTTATTGCAA ATCAAGATTT TATTAAGAAA CTCACCTTTA	3325
CACAATATAA TAAGATTGCC TATATTGACA TGGACATAGA AACCACAGAA TTTAAGATAT	3385
TAAGATTAGT AGTGTGTAAG ACTAGAATAA ATATTTATGT TTGCAACGTA ATTGGTAAAT	3445
TGAAAGAAAC TAATTTTGAA AA	3467

Sequence Number : 4

Sequence Length : 680

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Protein

Original Source

Organism : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

```
Met Ala Ser Pro Gly Ser Asn His Gly Tyr Asp Val Ile Asp His Ser
 1           5           10           15
Arg Ile Asn Asp Glu Leu Gly Gly Glu Lys Glu Tyr Arg Arg Leu Ile
          20           25           30
Glu Thr Ala His Thr Ile Gly Leu Gly Ile Ile Gln Asp Ile Val Pro
          35           40           45
Asn His Met Ala Val Asn Ser Leu Asn Trp Arg Leu Met Asp Val Leu
          50           55           60
Lys Met Gly Lys Lys Ser Lys Tyr Tyr Thr Tyr Phe Asp Phe Phe Pro
          65           70           75           80
Glu Asp Asp Lys Ile Arg Leu Pro Ile Leu Gly Glu Asp Leu Asp Thr
          85           90           95
Val Ile Ser Lys Gly Leu Leu Lys Ile Val Lys Asp Gly Asp Glu Tyr
          100          105          110
```

Phe Leu Glu Tyr Phe Lys Trp Lys Leu Pro Leu Thr Glu Val Gly Asn

115

120

125

Asp Ile Tyr Asp Thr Leu Gln Lys Gln Asn Tyr Thr Leu Met Ser Trp

130

135

140

Lys Asn Pro Pro Ser Tyr Arg Arg Phe Phe Asp Val Asn Thr Leu Ile

145

150

155

160

Gly Val Asn Val Glu Lys Asp His Val Phe Gln Glu Ser His Ser Lys

165

170

175

Ile Leu Asp Leu Asp Val Asp Gly Tyr Arg Ile Asp His Ile Asp Gly

180

185

190

Leu Tyr Asp Pro Glu Lys Tyr Ile Asn Asp Leu Arg Ser Ile Ile Lys

195

200

205

Asn Lys Ile Ile Ile Val Glu Lys Ile Leu Gly Phe Gln Glu Glu Leu

210

215

220

Lys Leu Asn Ser Asp Gly Thr Thr Gly Tyr Asp Phe Leu Asn Tyr Ser

225

230

235

240

Asn Leu Leu Phe Asn Phe Asn Gln Glu Ile Met Asp Ser Ile Tyr Glu

245

250

255

Asn Phe Thr Ala Glu Lys Ile Ser Ile Ser Glu Ser Ile Lys Lys Ile

260

265

270

Lys Ala Gln Ile Ile Asp Glu Leu Phe Ser Tyr Glu Val Lys Arg Leu

275

280

285

Ala Ser Gln Leu Gly Ile Ser Tyr Asp Ile Leu Arg Asp Tyr Leu Ser

290

295

300

Cys Ile Asp Val Tyr Arg Thr Tyr Ala Asn Gln Ile Val Lys Glu Cys
305 310 315 320
Asp Lys Thr Asn Glu Ile Glu Glu Ala Thr Lys Arg Asn Pro Glu Ala
325 330 335
Tyr Thr Lys Leu Gln Gln Tyr Met Pro Ala Val Tyr Ala Lys Ala Tyr
340 345 350
Glu Asp Thr Phe Leu Phe Arg Tyr Asn Arg Leu Ile Ser Ile Asn Glu
355 360 365
Val Gly Ser Asp Leu Arg Tyr Tyr Lys Ile Ser Pro Asp Gln Phe His
370 375 380
Val Phe Asn Gln Lys Arg Arg Gly Lys Ile Thr Leu Asn Ala Thr Ser
385 390 395 400
Thr His Asp Thr Lys Phe Ser Glu Asp Val Arg Met Lys Ile Ser Val
405 410 415
Leu Ser Glu Phe Pro Glu Glu Trp Lys Asn Lys Val Glu Glu Trp His
420 425 430
Ser Ile Ile Asn Pro Lys Val Ser Arg Asn Asp Glu Tyr Arg Tyr Tyr
435 440 445
Gln Val Leu Val Gly Ser Phe Tyr Glu Gly Phe Ser Asn Asp Phe Lys
450 455 460
Glu Arg Ile Lys Gln His Met Ile Lys Ser Val Arg Glu Ala Lys Ile
465 470 475 480
Asn Thr Ser Trp Arg Asn Gln Asn Lys Glu Tyr Glu Asn Arg Val Met
485 490 495

Glu Leu Val Glu Glu Thr Phe Thr Asn Lys Asp Phe Ile Lys Ser Phe

500

505

510

Met Lys Phe Glu Ser Lys Ile Arg Arg Ile Gly Met Ile Lys Ser Leu

515

520

525

Ser Leu Val Ala Leu Lys Ile Met Ser Ala Gly Ile Pro Asp Phe Tyr

530

535

540

Gln Gly Thr Glu Ile Trp Arg Tyr Leu Leu Thr Asp Pro Asp Asn Arg

545

550

555

560

Val Pro Val Asp Phe Lys Lys Leu His Glu Ile Leu Glu Lys Ser Lys

565

570

575

Lys Phe Glu Lys Asn Met Leu Glu Ser Met Asp Asp Gly Arg Ile Lys

580

585

590

Met Tyr Leu Thr Tyr Lys Leu Leu Ser Leu Arg Lys Gln Leu Ala Glu

595

600

605

Asp Phe Leu Lys Gly Glu Tyr Lys Gly Leu Asp Leu Glu Glu Gly Leu

610

615

620

Cys Gly Phe Ile Arg Phe Asn Lys Ile Leu Val Ile Ile Lys Thr Lys

625

630

635

640

Gly Ser Val Asn Tyr Lys Leu Lys Leu Glu Glu Gly Ala Ile Tyr Thr

645

650

655

Asp Val Leu Thr Gly Glu Glu Ile Lys Lys Glu Val Gln Ile Asn Glu

660

665

670

Leu Pro Arg Ile Leu Val Arg Met

675

680

Sequence Number : 5

Sequence Length : 2691

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Genomic DNA

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

```
CTGCAGTAAC TAGCGCTATC GAAGACGTTA TAAAGAGAAG GATAAATAGA GTTCCAGTGA   60
GTCTAGAAGA CCTTTTGTAA TAAGGACTTT AATATCATTT AAATTTATTT TTTGGAACAT  120
GCAGAGGTAA ACCCATGAAT GTCATTTTCG ACGTATTTAA CGAGATCCAT GGGTTTTTTG  180
GTGCATTGTG GCGGGGAGCA GCTCTACTTA ACTACTTAGT TAAGCCTCAA GATAAGAGGC  240
AATTTGAGAG AATAGGGAAA TTCTTCATGA TAAACTCAGT CATTACAGTA ATAACTGGGA  300
TAATAATTTT CGCCTACATT TACCTAGCCC CTTATCAAGG GAATTTATTT CTAGTAGCGG  360
CAATTCTACG TTCAAGCCTT GACATTAGGT TAAGGGCCTT ACTAAACTTA ATAGGAGGAG  420
CGTTTGGGTT ATTGGCTTTT GGGGCAGGGA TAGTTATAAG CAATAGGATA AGGCTTATGG  480
TACGTGTTAA GGAAGGTGAC GCTACAATCC TAGAGTTGAG GAATAGTATT GCCAATTTAT  540
CTAAAATTAG TTTAATCTTC TTATTACTTT CCTTAGCCAT GATGATACTT GCTGGTTCCA  600
TAGCACAAGT TATAAGTTAG AGTTGAAAGA AAAATTTA ATG ACG TTT GCT TAT AAA  656
```

Met Thr Phe Ala Tyr Lys

0 1

5

ATA GAT GGA AAT GAG GTA ATC TTT ACC TTA TGG GCA CCT TAT CAA AAG	704
Ile Asp Gly Asn Glu Val Ile Phe Thr Leu Trp Ala Pro Tyr Gln Lys	
10 15 20	
AGC GTT AAA CTA AAG GTT CTA GAG AAG GGA CTT TAC GAA ATG GAA AGA	752
Ser Val Lys Leu Lys Val Leu Glu Lys Gly Leu Tyr Glu Met Glu Arg	
25 30 35	
GAT GAA AAA GGT TAC TTC ACC ATT ACC TTA AAC AAC GTA AAG GTT AGA	800
Asp Glu Lys Gly Tyr Phe Thr Ile Thr Leu Asn Asn Val Lys Val Arg	
40 45 50	
GAT AGG TAT AAA TAC GTT TTA GAT GAT GCT AGT GAA ATA CCA GAT CCA	848
Asp Arg Tyr Lys Tyr Val Leu Asp Asp Ala Ser Glu Ile Pro Asp Pro	
55 60 65	
GCA TCC AGA TAC CAA CCA GAA GGT GTA CAT GGG CCT TCA CAA ATT ATA	896
Ala Ser Arg Tyr Gln Pro Glu Gly Val His Gly Pro Ser Gln Ile Ile	
70 75 80 85	
CAA GAA AGT AAA GAG TTC AAC AAC GAG ACT TTT CTG AAG AAA GAG GAC	944
Gln Glu Ser Lys Glu Phe Asn Asn Glu Thr Phe Leu Lys Lys Glu Asp	
90 95 100	
TTG ATA ATT TAT GAA ATA CAC GTG GGG ACT TTC ACT CCA GAG GGA ACG	992
Leu Ile Ile Tyr Glu Ile His Val Gly Thr Phe Thr Pro Glu Gly Thr	
105 110 115	
TTT GAG GGA GTG ATA AGG AAA CTT GAC TAC TTA AAG GAT TTG GGA ATT	1040
Phe Glu Gly Val Ile Arg Lys Leu Asp Tyr Leu Lys Asp Leu Gly Ile	
120 125 130	

ACG GCA ATA GAG ATA ATG CCA ATA GCT CAA TTT CCT GGG AAA AGG GAT 1088

Thr Ala Ile Glu Ile Met Pro Ile Ala Gln Phe Pro Gly Lys Arg Asp

135

140

145

TGG GGT TAT GAT GGA GTT TAT TTA TAT GCA GTA CAG AAC TCT TAC GGA 1136

Trp Gly Tyr Asp Gly Val Tyr Leu Tyr Ala Val Gln Asn Ser Tyr Gly

150

155

160

165

GGG CCA GAA GGT TTT AGA AAG TTA GTT GAT GAA GCG CAC AAG AAA GGT 1184

Gly Pro Glu Gly Phe Arg Lys Leu Val Asp Glu Ala His Lys Lys Gly

170

175

180

TTA GGA GTT ATT TTA GAC GTA GTA TAC AAC CAC GTT GGA CCA GAG GGA 1232

Leu Gly Val Ile Leu Asp Val Val Tyr Asn His Val Gly Pro Glu Gly

185

190

195

AAC TAT ATG GTT AAA TTG GGG CCA TAT TTC TCA CAG AAA TAC AAA ACG 1280

Asn Tyr Met Val Lys Leu Gly Pro Tyr Phe Ser Gln Lys Tyr Lys Thr

200

205

210

CCA TGG GGA TTA ACC TTT AAC TTT GAC GAT GCT GAA AGC GAT GAG GTT 1328

Pro Trp Gly Leu Thr Phe Asn Phe Asp Asp Ala Glu Ser Asp Glu Val

215

220

225

AGG AAG TTC ATC TTA GAA AAC GTT GAG TAC TGG ATT AAG GAA TAT AAC 1376

Arg Lys Phe Ile Leu Glu Asn Val Glu Tyr Trp Ile Lys Glu Tyr Asn

230

235

240

245

GTT GAT GGG TTT AGA TTA GAT GCG GTT CAT GCA ATT ATT GAC ACT TCT 1424

Val Asp Gly Phe Arg Leu Asp Ala Val His Ala Ile Ile Asp Thr Ser

250

255

260

CCT AAG CAC ATC TTG GAG GAA ATA GCT GAC GTT GTG CAT AAG TAT AAT 1472

Pro Lys His Ile Leu Glu Glu Ile Ala Asp Val Val His Lys Tyr Asn

265

270

275

AGG ATT GTC ATA GCC GAA AGT GAT TTA AAC GAT CCT AGA GTC GTT AAT 1520

Arg Ile Val Ile Ala Glu Ser Asp Leu Asn Asp Pro Arg Val Val Asn

280

285

290

CCC AAG GAA AAG TGT GGA TAT AAT ATT GAT GCT CAA TGG GTT GAC GAT 1568

Pro Lys Glu Lys Cys Gly Tyr Asn Ile Asp Ala Gln Trp Val Asp Asp

295

300

305

TTC CAT CAT TCT ATT CAC GCT TAC TTA ACT GGT GAG AGG CAA GGC TAT 1616

Phe His His Ser Ile His Ala Tyr Leu Thr Gly Glu Arg Gln Gly Tyr

310

315

320

325

TAT ACG GAT TTC GGT AAC CTT GAC GAT ATA GTT AAA TCG TAT AAG GAC 1664

Tyr Thr Asp Phe Gly Asn Leu Asp Asp Ile Val Lys Ser Tyr Lys Asp

330

335

340

GTT TTC GTA TAT GAT GGT AAG TAC TCC AAT TTT AGA AGA AAA ACT CAC 1712

Val Phe Val Tyr Asp Gly Lys Tyr Ser Asn Phe Arg Arg Lys Thr His

345

350

355

GGA GAA CCA GTT GGT GAA CTA GAC GGA TGC AAT TTC GTA GTT TAT ATA 1760

Gly Glu Pro Val Gly Glu Leu Asp Gly Cys Asn Phe Val Val Tyr Ile

360

365

370

CAA AAT CAC GAT CAA GTC GGA AAT AGA GGC AAA GGT GAA AGA ATA ATT 1808

Gln Asn His Asp Gln Val Gly Asn Arg Gly Lys Gly Glu Arg Ile Ile

375

380

385

AAA TTA GTC GAT AGG GAA AGC TAC AAG ATC GCT GCA GCC CTT TAC CTT	1856
Lys Leu Val Asp Arg Glu Ser Tyr Lys Ile Ala Ala Ala Leu Tyr Leu	
390 395 400 405	
CTT TCC CCC TAT ATT CCA ATG ATT TTC ATG GGA GAG GAA TAC GGT GAG	1904
Leu Ser Pro Tyr Ile Pro Met Ile Phe Met Gly Glu Glu Tyr Gly Glu	
410 415 420	
GAA AAT CCC TTT TAT TTC TTT TCT GAT TTT TCA GAT TCA AAA CTG ATA	1952
Glu Asn Pro Phe Tyr Phe Phe Ser Asp Phe Ser Asp Ser Lys Leu Ile	
425 430 435	
CAA GGT GTA AGG GAA GGG AGA AAA AAG GAA AAC GGG CAA GAT ACT GAC	2000
Gln Gly Val Arg Glu Gly Arg Lys Lys Glu Asn Gly Gln Asp Thr Asp	
440 445 450	
CCT CAA GAT GAA TCA ACT TTT AAC GCT TCC AAA CTG AGT TGG AAG ATT	2048
Pro Gln Asp Glu Ser Thr Phe Asn Ala Ser Lys Leu Ser Trp Lys Ile	
455 460 465	
GAC GAG GAA ATC TTT TCA TTT TAC AAG ATT TTA ATA AAA ATG AGA AAG	2096
Asp Glu Glu Ile Phe Ser Phe Tyr Lys Ile Leu Ile Lys Met Arg Lys	
470 475 480 485	
GAG TTG AGC ATA GCG TGT GAT AGG AGA GTA AAC GTC GTG AAT GGC GAA	2144
Glu Leu Ser Ile Ala Cys Asp Arg Arg Val Asn Val Val Asn Gly Glu	
490 495 500	
AAT TGC TTG ATC ATC AAG GCA AGA GAA TAC TTT TCA CTC TAC GTT TTC	2192
Asn Trp Leu Ile Ile Lys Gly Arg Glu Tyr Phe Ser Leu Tyr Val Phe	
505 510 515	

TCT AAA TCA TCT ATT GAA GTT AAG TAC AGT GGA ACT TTA CTT TTG TCC 2240
Ser Lys Ser Ser Ile Glu Val Lys Tyr Ser Gly Thr Leu Leu Leu Ser
520 525 530
TCA AAT AAT TCA TTC CCT CAG CAT ATT GAA GAA GGT AAA TAT GAG TTT 2288
Ser Asn Asn Ser Phe Pro Gln His Ile Glu Glu Gly Lys Tyr Glu Phe
535 540 545
GAT AAG GGA TTT GCT TTA TAT AAA CTT TAGGACA GGAGAGTTTA AAAATTTCTA 2342
Asp Lys Gly Phe Ala Leu Tyr Lys Leu
550 555
TGAATGATTA TACTTTAGAT GATGAGTAAA AGCAAGATCG ATGAGGAAGA GAAAAGGAGA 2402
AGAGAAGAAG TCAAAAAGTT AGTAATGCTC TTAGCAATGT TAAGATAATG TTTTTTTAAA 2462
CTCAAATAAT AATAAATACC ATCATGTCAA TATTCTTCAG AACTAGAGAT AGACCTTTAC 2522
GTCCCGGAGA TCCGTATCCA TTAGGTTCAA ATTGGATAGA AGATGAGGAT GCGGTAAATT 2582
TTTCCTTGTT CTCAGAGAAT GCAGACAAAG TGGAGTTGAT TCTTTATTCA CAAACAAATC 2642
AAAAGTATCC AAAGGAGATA ATAGAGGTTA AGAATAGAAC GGGGGATCC 2691

Sequence Number : 6

Sequence Length : 558

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Protein

Original Source

Organism : Sulfolobus solfataricus

Strain : KM1

Sequence

Thr Phe Ala Tyr Lys Ile Asp Gly Asn Glu Val Ile Phe Thr Leu Trp
1 5 10 15
Ala Pro Tyr Gln Lys Ser Val Lys Leu Lys Val Leu Glu Lys Gly Leu
20 25 30
Tyr Glu Met Glu Arg Asp Glu Lys Gly Tyr Phe Thr Ile Thr Leu Asn
35 40 45
Asn Val Lys Val Arg Asp Arg Tyr Lys Tyr Val Leu Asp Asp Ala Ser
50 55 60
Glu Ile Pro Asp Pro Ala Ser Arg Tyr Gln Pro Glu Gly Val His Gly
65 70 75 80
Pro Ser Gln Ile Ile Gln Glu Ser Lys Glu Phe Asn Asn Glu Thr Phe
85 90 95
Leu Lys Lys Glu Asp Leu Ile Ile Tyr Glu Ile His Val Gly Thr Phe
100 105 110
Thr Pro Glu Gly Thr Phe Glu Gly Val Ile Arg Lys Leu Asp Tyr Leu
115 120 125
Lys Asp Leu Gly Ile Thr Ala Ile Glu Ile Met Pro Ile Ala Gln Phe
130 135 140
Pro Gly Lys Arg Asp Trp Gly Tyr Asp Gly Val Tyr Leu Tyr Ala Val
145 150 155 160
Gln Asn Ser Tyr Gly Gly Pro Glu Gly Phe Arg Lys Leu Val Asp Glu
165 170 175

Ala His Lys Lys Gly Leu Gly Val Ile Leu Asp Val Val Tyr Asn His

180

185

190

Val Gly Pro Glu Gly Asn Tyr Met Val Lys Leu Gly Pro Tyr Phe Ser

195

200

205

Gln Lys Tyr Lys Thr Pro Trp Gly Leu Thr Phe Asn Phe Asp Asp Ala

210

215

220

Glu Ser Asp Glu Val Arg Lys Phe Ile Leu Glu Asn Val Glu Tyr Trp

225

230

235

240

Ile Lys Glu Tyr Asn Val Asp Gly Phe Arg Leu Asp Ala Val His Ala

245

250

255

Ile Ile Asp Thr Ser Pro Lys His Ile Leu Glu Glu Ile Ala Asp Val

260

265

270

Val His Lys Tyr Asn Arg Ile Val Ile Ala Glu Ser Asp Leu Asn Asp

275

280

285

Pro Arg Val Val Asn Pro Lys Glu Lys Cys Gly Tyr Asn Ile Asp Ala

290

295

300

Gln Trp Val Asp Asp Phe His His Ser Ile His Ala Tyr Leu Thr Gly

305

310

315

320

Glu Arg Gln Gly Tyr Tyr Thr Asp Phe Gly Asn Leu Asp Asp Ile Val

325

330

335

Lys Ser Tyr Lys Asp Val Phe Val Tyr Asp Gly Lys Tyr Ser Asn Phe

340

345

350

Arg Arg Lys Thr His Gly Glu Pro Val Gly Glu Leu Asp Gly Cys Asn

355

360

365

Phe Val Val Tyr Ile Gln Asn His Asp Gln Val Gly Asn Arg Gly Lys
370 375 380
Gly Glu Arg Ile Ile Lys Leu Val Asp Arg Glu Ser Tyr Lys Ile Ala
385 390 395 400
Ala Ala Leu Tyr Leu Leu Ser Pro Tyr Ile Pro Met Ile Phe Met Gly
405 410 415
Glu Glu Tyr Gly Glu Glu Asn Pro Phe Tyr Phe Phe Ser Asp Phe Ser
420 425 430
Asp Ser Lys Leu Ile Gln Gly Val Arg Glu Gly Arg Lys Lys Glu Asn
435 440 445
Gly Gln Asp Thr Asp Pro Gln Asp Glu Ser Thr Phe Asn Ala Ser Lys
450 455 460
Leu Ser Trp Lys Ile Asp Glu Glu Ile Phe Ser Phe Tyr Lys Ile Leu
465 470 475 480
Ile Lys Met Arg Lys Glu Leu Ser Ile Ala Cys Asp Arg Arg Val Asn
485 490 495
Val Val Asn Gly Glu Asn Trp Leu Ile Ile Lys Gly Arg Glu Tyr Phe
500 505 510
Ser Leu Tyr Val Phe Ser Lys Ser Ser Ile Glu Val Lys Tyr Ser Gly
515 520 525
Thr Leu Leu Leu Ser Ser Asn Asn Ser Phe Pro Gln His Ile Glu Glu
530 535 540
Gly Lys Tyr Glu Phe Asp Lys Gly Phe Ala Leu Tyr Lys Leu
545 550 555

Sequence Number : 7

Sequence Length : 3600

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Genomic DNA

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

ATTCGTTTTG AGTCACTCGG CGTAGGTCTG TAGTCTTTCT TGGCGAGGGC TAATAAGTTG	60
AGATAATGCT TGCCAAGAAT CGAAGAAGGC GTCCTGCCCT GCATGAAATC GATTACCTCG	120
GCACTAACTC CGAGCTCCGC GAGTTTAGTA GTCACGAATT TCCGTACATA TTTCGGCGCT	180
ATCCCTTTCT CATGCAATAA ATTCTTCGGC TAGTTGTACG TTATATCAGT CTTAGCTATA	240
GACGAAATGT GAAAGACATA GAACACTTTC TTTGGCCCTC TAGTCCAGTT GAGCGTGTAT	300
ACGTAGAAGC CGTCCTCTTT CACGTTGTTC TTCTCGTCAT ACTCATTGAG AACCTTTACA	360
GCCTCCCTAA GCCTTATACC GCTCTCAAGG AGGAGCTTGA AGACTAGCTC TACCTCAATA	420
CCTCTAACAG CCTCCAACCA CCTCCCTATC TCGTCAGCTC CTGGAACCTT AAGATCAACA	480
CCAGACTTTT TCGTTTTCAG CTTTTTCCAT GCCTCAAGAT CCCCTTTCCA CTTGTAGAAC	540
TTCTTCCAGG CTAGGATAGA GTTCTTAGCA TTAGTAGGGG GCTTCTTCAG ATAATTGATA	600
TACTGCCTGC AAGTTTCCTC ACTGGCCATT TTCAAACAAT ATTCATAAAA TTCAATTAAT	660
TCCTTTTCCG TGAGACCATT TTTGCCCTCC CTAGAAGTAA GGGAGTTTAG GGCAAATCCC	720
TTACTCTCTT CATCATTTGA AAGAGGGGTT TTAGGGGATT CCTCCCCTAA CCAGGGCTTT	780
GGCCCCCTGG ACCAGGGTTC GAGTCCCTGC CCGGCTACCT TTGAAAGGTT AGGGGGATAC	840
ACCCTAATAC CCCACTTCTA TCTTACAATT TCAGGTAAGT CTTTACTAGG TCAACTAAAG	900

CACCAACGTA AGTCTCCTTC GTCTTACCAC CTTGACTCTT CTTGATAAAG TAAACATAAT 960
 ATCATCCATA GACTTACCTT ATTCTTATAT TACCATATGA TTTTATTATT TTGTATTTCT 1020
 ATTAGATAAG TCCCACATCAT AGAACAAATG ATGGTTTTAA CTTATATACT AAATACTCTA 1080
 ATAACCTCAAC AATAATAAGA ATTTAATCAG TTCTGATAAG TATTTTCACT CGAAAACATT 1140
 TAAATATATT AAGACATAAT TTCTATTTAA ACAGC ATG TTT TCG TTC GGT GGA AAT 1196

Met Phe Ser Phe Gly Gly Asn

1 5

ATT GAA AAA AAT AAA GGT ATC TTT AAG TTA TGG GCA CCT TAT GTT AAT 1244
 Ile Glu Lys Asn Lys Gly Ile Phe Lys Leu Trp Ala Pro Tyr Val Asn

10 15 20

AGT GTT AAG CTG AAG TTA AGC AAA AAA CTT ATT CCA ATG GAA AAA AAC 1292
 Ser Val Lys Leu Lys Leu Ser Lys Lys Leu Ile Pro Met Glu Lys Asn

25 30 35

GAT GAG GGA TTT TTC GAA GTA GAA ATA GAC GAT ATC GAG GAA AAT TTA 1340
 Asp Glu Gly Phe Phe Glu Val Glu Ile Asp Asp Ile Glu Glu Asn Leu

40 45 50 55

ACC TAT TCT TAT ATT ATA GAA GAT AAG AGA GAG ATA CCT GAT CCC GCA 1388
 Thr Tyr Ser Tyr Ile Ile Glu Asp Lys Arg Glu Ile Pro Asp Pro Ala

60 65 70

TCA CGA TAT CAA CCT TTA GGA GTT CAT GAC AAA TCA CAA CTT ATA AGA 1436
 Ser Arg Tyr Gln Pro Leu Gly Val His Asp Lys Ser Gln Leu Ile Arg

75 80 85

ACA GAT TAT CAG ATT CTT GAC CTT GGA AAA GTA AAA ATA GAA GAT CTA	1484
Thr Asp Tyr Gln Ile Leu Asp Leu Gly Lys Val Lys Ile Glu Asp Leu	
90 95 100	
ATA ATA TAT GAA CTC CAC GTT GGT ACT TTT TCC CAA GAA GGA AAT TTC	1532
Ile Ile Tyr Glu Leu His Val Gly Thr Phe Ser Gln Glu Gly Asn Phe	
105 110 115	
AAA GGA GTA ATA GAA AAG TTA GAT TAC CTC AAG GAT CTA GGA ATC ACA	1580
Lys Gly Val Ile Glu Lys Leu Asp Tyr Leu Lys Asp Leu Gly Ile Thr	
120 125 130 135	
GGA ATT GAA CTG ATG CCT GTG GCA CAA TTT CCA GCG AAT AGA GAT TGG	1628
Gly Ile Glu Leu Met Pro Val Ala Gln Phe Pro Gly Asn Arg Asp Trp	
140 145 150	
GGA TAC GAT GGT GTT TTT CTA TAC GCA GTT CAA AAT ACT TAT GGC GGA	1676
Gly Tyr Asp Gly Val Phe Leu Tyr Ala Val Gln Asn Thr Tyr Gly Gly	
155 160 165	
CCA TGG GAA TTG GCT AAG CTA GTA AAC GAG GCA CAT AAA AGG GGA ATA	1724
Pro Trp Glu Leu Ala Lys Leu Val Asn Glu Ala His Lys Arg Gly Ile	
170 175 180	
GCC GTA ATT TTG GAT GTT GTA TAT AAT CAT ATA GGT CCT GAG GGA AAT	1772
Ala Val Ile Leu Asp Val Val Tyr Asn His Ile Gly Pro Glu Gly Asn	
185 190 195	
TAC CTT TTA GGA TTA GGT CCT TAT TTT TCA GAC AGA TAT AAA ACT CCA	1820
Tyr Leu Leu Gly Leu Gly Pro Tyr Phe Ser Asp Arg Tyr Lys Thr Pro	
200 205 210 215	

TGG GGA TTA ACA TTT AAT TTT GAT GAT AGG GGA TGT GAT CAA GTT AGA	1868
Trp Gly Leu Thr Phe Asn Phe Asp Asp Arg Gly Cys Asp Gln Val Arg	
220 225 230	
AAA TTC ATT TTA GAA AAT GTC GAG TAT TGG TTT AAG ACC TTT AAA ATC	1916
Lys Phe Ile Leu Glu Asn Val Glu Tyr Trp Phe Lys Thr Phe Lys Ile	
235 240 245	
GAT GGT CTG AGA CTG GAT GCA GTT CAT GCA ATT TTT GAT AAT TCG CCT	1964
Asp Gly Leu Arg Leu Asp Ala Val His Ala Ile Phe Asp Asn Ser Pro	
250 255 260	
AAG CAT ATC CTC CAA GAG ATA GCT GAA AAA GCC CAT CAA TTA GGA AAA	2012
Lys His Ile Leu Gln Glu Ile Ala Glu Lys Ala His Gln Leu Gly Lys	
265 270 275	
TTT GTT ATT GCT GAA AGT GAT TTA AAT GAT CCA AAA ATA GTA AAA GAT	2060
Phe Val Ile Ala Glu Ser Asp Leu Asn Asp Pro Lys Ile Val Lys Asp	
280 285 290 295	
GAT TGT GGA TAT AAA ATA GAT GCT CAA TGG GTT GAC GAT TTC CAC CAC	2108
Asp Cys Gly Tyr Lys Ile Asp Ala Gln Trp Val Asp Asp Phe His His	
300 305 310	
GCA GTT CAT GCA TTC ATA ACA AAA GAA AAA GAT TAT TAT TAC CAG GAT	2156
Ala Val His Ala Phe Ile Thr Lys Glu Lys Asp Tyr Tyr Tyr Gln Asp	
315 320 325	
TTT GGA AGC ATA GAA GAT ATA GAG AAA ACT TTT AAA GAT GTT TTT GTT	2204
Phe Gly Arg Ile Glu Asp Ile Glu Lys Thr Phe Lys Asp Val Phe Val	
330 335 340	

TAT GAT GGA AAG TAT TCT AGA TAC AGA GGA AGA ACT CAT GGT GCT CCT	2252
Tyr Asp Gly Lys Tyr Ser Arg Tyr Arg Gly Arg Thr His Gly Ala Pro	
345 350 355	
GTA GGT GAT CTT CCA CCA CGT AAA TTT GTA GTC TTC ATA CAA AAT CAC	2300
Val Gly Asp Leu Pro Pro Arg Lys Phe Val Val Phe Ile Gln Asn His	
360 365 370 375	
GAT CAA GTA GGA AAT AGA GGA AAT GGG GAA AGA CTT TCC ATA TTA ACC	2348
Asp Gln Val Gly Asn Arg Gly Asn Gly Glu Arg Leu Ser Ile Leu Thr	
380 385 390	
GAT AAA ACG ACA TAC CTT ATG GCA GCC ACA CTA TAT ATA CTC TCA CCG	2396
Asp Lys Thr Thr Tyr Leu Met Ala Ala Thr Leu Tyr Ile Leu Ser Pro	
395 400 405	
TAT ATA CCG CTA ATA TTT ATG GGC GAG GAA TAT TAT GAG ACG AAT CCT	2444
Tyr Ile Pro Leu Ile Phe Met Gly Glu Glu Tyr Tyr Glu Thr Asn Pro	
410 415 420	
TTT TTC TTC TTC TCT GAT TTC TCA GAT CCC GTA TTA ATT AAG GGT GTT	2492
Phe Phe Phe Phe Ser Asp Phe Ser Asp Pro Val Leu Ile Lys Gly Val	
425 430 435	
AGA GAA GGT AGA CTA AAG GAA AAT AAT CAA ATG ATA GAT CCA CAA TCT	2540
Arg Glu Gly Arg Leu Lys Glu Asn Asn Gln Met Ile Asp Pro Gln Ser	
440 445 450 455	
GAG GAA GCG TTC TTA AAG AGT AAA CTT TCA TCG AAA ATT GAT GAG GAA	2588
Glu Glu Ala Phe Leu Lys Ser Lys Leu Ser Trp Lys Ile Asp Glu Glu	
460 465 470	

GTT TTA GAT TAT TAT AAA CAA CTG ATA AAT ATC AGA AAG AGA TAT AAT	2636
Val Leu Asp Tyr Tyr Lys Gln Leu Ile Asn Ile Arg Lys Arg Tyr Asn	
475 480 485	
AAT TGT AAA AGG GTA AAG GAA GTT AGG AGA GAA GGG AAC TGT ATT ACT	2684
Asn Cys Lys Arg Val Lys Glu Val Arg Arg Glu Gly Asn Cys Ile Thr	
490 495 500	
TTG ATC ATG GAA AAA ATA GGA ATA ATT GCA TCG TTT GAT GAT ATT GTA	2732
Leu Ile Met Glu Lys Ile Gly Ile Ile Ala Ser Phe Asp Asp Ile Val	
505 510 515	
ATT AAT TCT AAA ATT ACA GGT AAT TTA CTT ATA GGC ATA GGA TTT CCG	2780
Ile Asn Ser Lys Ile Thr Gly Asn Leu Leu Ile Gly Ile Gly Phe Pro	
520 525 530 535	
AAA AAA TTG AAA AAA GAT GAA TTA ATT AAG GTT AAC AGA GGT GTT GGG	2828
Lys Lys Leu Lys Lys Asp Glu Leu Ile Lys Val Asn Arg Gly Val Gly	
540 545 550	
GTA TAT CAA TTA GAA TGAAAGATCG ACCATTAAAG CCTGGTGAAC CTTATCCTTT	2883
Val Tyr Gln Leu Glu	
555	
AGGGGCAACT TGGATAGAGG AAGAAGATGG AGTTAATTTT GTACTATTCT CTGAGAACGC	2943
CACAAAAGTA GAACTGTTAA CGTACTCTCA GACTAGACAA GATGAGCCAA AGGAAATAAT	3003
AGAACTTAGA CAGAGAACCG GAGATCTCTG GCATGTTTTT GTACCTGGTT TAAGACCAGG	3063
TCAGTTGTAT GGGTACAGGG TGTATGGTCC ATATAAACCA GAGGAAGGGT TAAGGTTTAA	3123
TCCTAATAAA GTACTGATAG ATCCTTATGC AAAAGCTATA AACGGATTAT TACTATGGGA	3183
TGATTCCGTC TTTGGATATA AAATTGGAGA TCAGAACCAG GATCTCAGTT TCGATGAGAG	3243

AAAAGACGAT AAATTTATAC CTAAAGGGGT CATAATAAAT CCTTATTTTG ATTGGGAGGA 3303
CGAGCATTTTC TTCTTTAGAA GAAAGATACC TTTTAAGGAT AGTATAATTT ATGAGACACA 3363
TATAAAAGGA ATAATAAAT TAAGGCAAGA TTTACCGGAG AACGTTAGAG GCACTTTTTT 3423
GGGTTTAGCA TCAGATACTA TGATTGATTA CCTAAAAGAT TTAGGAATTA CAACCGTTGA 3483
GATAATGCCT ATTCAGCAAT TTGTAGATGA GAGATTCATT GTCGATAAAG GGTAAAGAA 3543
CTACTGGGGT TACAATCCGA TAAATTATTT CTCTCCTGAA TGTAGATACT CAAGCTC 3600

Sequence Number : 8

Sequence Length : 556

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Protein

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

Met Phe Ser Phe Gly Gly Asn Ile Glu Lys Asn Lys Gly Ile Phe Lys

1 5 10 15

Leu Trp Ala Pro Tyr Val Asn Ser Val Lys Leu Lys Leu Ser Lys Lys

20 25 30

Leu Ile Pro Met Glu Lys Asn Asp Glu Gly Phe Phe Glu Val Glu Ile

35 40 45

Asp Asp Ile Glu Glu Asn Leu Thr Tyr Ser Tyr Ile Ile Glu Asp Lys

50 55 60

Arg	Glu	Ile	Pro	Asp	Pro	Ala	Ser	Arg	Tyr	Gln	Pro	Leu	Gly	Val	His
65					70					75					80
Asp	Lys	Ser	Gln	Leu	Ile	Arg	Thr	Asp	Tyr	Gln	Ile	Leu	Asp	Leu	Gly
			85						90					95	
Lys	Val	Lys	Ile	Glu	Asp	Leu	Ile	Ile	Tyr	Glu	Leu	His	Val	Gly	Thr
			100					105						110	
Phe	Ser	Gln	Glu	Gly	Asn	Phe	Lys	Gly	Val	Ile	Glu	Lys	Leu	Asp	Tyr
			115					120						125	
Leu	Lys	Asp	Leu	Gly	Ile	Thr	Gly	Ile	Glu	Leu	Met	Pro	Val	Ala	Gln
		130					135						140		
Phe	Pro	Gly	Asn	Arg	Asp	Trp	Gly	Tyr	Asp	Gly	Val	Phe	Leu	Tyr	Ala
145					150					155				160	
Val	Gln	Asn	Thr	Tyr	Gly	Gly	Pro	Trp	Glu	Leu	Ala	Lys	Leu	Val	Asn
					165				170					175	
Glu	Ala	His	Lys	Arg	Gly	Ile	Ala	Val	Ile	Leu	Asp	Val	Val	Tyr	Asn
			180					185						190	
His	Ile	Gly	Pro	Glu	Gly	Asn	Tyr	Leu	Leu	Gly	Leu	Gly	Pro	Tyr	Phe
			195					200						205	
Ser	Asp	Arg	Tyr	Lys	Thr	Pro	Trp	Gly	Leu	Thr	Phe	Asn	Phe	Asp	Asp
			210					215						220	
Arg	Gly	Cys	Asp	Gln	Val	Arg	Lys	Phe	Ile	Leu	Glu	Asn	Val	Glu	Tyr
225					230					235				240	
Trp	Phe	Lys	Thr	Phe	Lys	Ile	Asp	Gly	Leu	Arg	Leu	Asp	Ala	Val	His
					245					250				255	

Ala Ile Phe Asp Asn Ser Pro Lys His Ile Leu Gln Glu Ile Ala Glu

260

265

270

Lys Ala His Gln Leu Gly Lys Phe Val Ile Ala Glu Ser Asp Leu Asn

275

280

285

Asp Pro Lys Ile Val Lys Asp Asp Cys Gly Tyr Lys Ile Asp Ala Gln

290

295

300

Trp Val Asp Asp Phe His His Ala Val His Ala Phe Ile Thr Lys Glu

305

310

315

320

Lys Asp Tyr Tyr Tyr Gln Asp Phe Gly Arg Ile Glu Asp Ile Glu Lys

325

330

335

Thr Phe Lys Asp Val Phe Val Tyr Asp Gly Lys Tyr Ser Arg Tyr Arg

340

345

350

Gly Arg Thr His Gly Ala Pro Val Gly Asp Leu Pro Pro Arg Lys Phe

355

360

365

Val Val Phe Ile Gln Asn His Asp Gln Val Gly Asn Arg Gly Asn Gly

370

375

380

Glu Arg Leu Ser Ile Leu Thr Asp Lys Thr Thr Tyr Leu Met Ala Ala

385

390

395

400

Thr Leu Tyr Ile Leu Ser Pro Tyr Ile Pro Leu Ile Phe Met Gly Glu

405

410

415

Glu Tyr Tyr Glu Thr Asn Pro Phe Phe Phe Phe Ser Asp Phe Ser Asp

420

425

430

Pro Val Leu Ile Lys Gly Val Arg Glu Gly Arg Leu Lys Glu Asn Asn

435

440

445

Gln Met Ile Asp Pro Gln Ser Glu Glu Ala Phe Leu Lys Ser Lys Leu
450 455 460
Ser Trp Lys Ile Asp Glu Glu Val Leu Asp Tyr Tyr Lys Gln Leu Ile
465 470 475 480
Asn Ile Arg Lys Arg Tyr Asn Asn Cys Lys Arg Val Lys Glu Val Arg
485 490 495
Arg Glu Gly Asn Cys Ile Thr Leu Ile Met Glu Lys Ile Gly Ile Ile
500 505 510
Ala Ser Phe Asp Asp Ile Val Ile Asn Ser Lys Ile Thr Gly Asn Leu
515 520 525
Leu Ile Gly Ile Gly Phe Pro Lys Lys Leu Lys Lys Asp Glu Leu Ile
530 535 540
Lys Val Asn Arg Gly Val Gly Val Tyr Gln Leu Glu
545 550 555

Sequence Number : 9

Sequence Length : 6

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : Sulfolobus solfataricus

Strain : KM1

Sequence

Val Ile Arg Glu Ala Lys

1

5

Sequence Number : 10

Sequence Length : 6

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Ile Ser Ile Arg Gln Lys

1

5

Sequence Number : 11

Sequence Length : 5

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Ile Ile Tyr Val Glu

1

5

Sequence Number : 12

Sequence Length : 5

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Met Leu Tyr Val Lys

1

5

Sequence Number : 13

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Ile Leu Ser Ile Asn Glu Lys

1

5

Sequence Number : 14

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Val Val Ile Leu Thr Glu Lys

1

5

Sequence Number : 15

Sequence Length : 10

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : Sulfolobus solfataricus

Strain : KM1

Sequence

Asn Leu Glu Leu Ser Asp Pro Arg Val Lys

1

5

10

Sequence Number : 16

Sequence Length : 12

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Met Ile Ile Gly Thr Tyr Arg Leu Gln Leu Asn Lys

1 5 10

Sequence Number : 17

Sequence Length : 9

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Val Ala Val Leu Phe Ser Pro Ile Val

1 5 9

Sequence Number : 18

Sequence Length : 11

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Ile Asn Ile Asp Glu Leu Ile Ile Gln Ser Lys

1

5

10

Sequence Number : 19

Sequence Length : 12

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Glu Leu Gly Val Ser His Leu Tyr Leu Ser Pro Ile

1

5

10

Sequence Number : 20

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Glu Val Phe Arg Glu Ser

1

5

Sequence Number : 21

Sequence Length : 4

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Tyr Phe Lys

1

Sequence Number : 22

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Gly Leu Tyr Asn Pro Lys

1

5

Sequence Number : 23

Sequence Length : 8

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Ile Asn Gly Ile Arg Glu Cys

1

5

Sequence Number : 24

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Phe Glu Asn Phe Glu Lys

1

5

Sequence Number : 25

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Leu Leu Arg Pro Asn Ile

1

5

Sequence Number : 26

Sequence Length : 5

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : Sulfolobus solfataricus

Strain : KM1

Sequence

Asp Ile Ile Glu Asn

1

5

Sequence Number : 27

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : Sulfolobus solfataricus

Strain : KM1

Sequence

Asp Asn Ile Glu Tyr Arg Gly

1

5

Sequence Number : 28

Sequence Length : 18

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Other nucleic acid (Synthesized DNA)

Sequence

YTCWCKRAAW ACYTCATC

Sequence Number : 29

Sequence Length : 20

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Other nucleic acid (Synthesized DNA)

Sequence

GATAAYATWG ARTAYAGRGG

Sequence Number : 30

Sequence Length : 8

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Arg Asn Pro Glu Ala Tyr Thr Lys

1

5

Sequence Number : 31

Sequence Length : 9

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Asp His Val Phe Gln Glu Ser His Ser

1

5

Sequence Number : 32

Sequence Length : 8

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Ile Thr Leu Asn Ala Thr Ser Thr

1

5

Sequence Number : 33

Sequence Length : 6

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Ile Ile Ile Val Glu Lys

1

5

Sequence Number : 34

Sequence Length : 11

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Leu Gln Gln Tyr Met Pro Ala Val Tyr Ala Lys

1

5

10

Sequence Number : 35

Sequence Length : 5

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Asn Met Leu Glu Ser

1 5

Sequence Number : 36

Sequence Length : 13

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

Lys Ile Ser Pro Asp Gln Phe His Val Phe Asn Gln Lys

1 5 10

Sequence Number : 37

Sequence Length : 8

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

Gln Leu Ala Glu Asp Phe Leu Lys

1

5

Sequence Number : 38

Sequence Length : 10

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Lys Ile Leu Gly Phe Gln Glu Glu Leu Lys

1

5

10

Sequence Number : 39

Sequence Length : 10

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Ile Ser Val Leu Ser Glu Phe Pro Glu Glu

1

5

10

Sequence Number : 40

Sequence Length : 9

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Leu Lys Leu Glu Glu Gly Ala Ile Tyr

1

5

Sequence Number : 41

Sequence Length : 8

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Glu Val Gln Ile Asn Glu Leu Pro

1

5

Sequence Number : 42

Sequence Length : 5

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Asp His Ser Arg Ile

1

5

Sequence Number : 43

Sequence Length : 6

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

Asp Leu Arg Tyr Tyr Lys

1

5

Sequence Number : 44

Sequence Length : 14

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

Asp Val Tyr Arg Thr Tyr Ala Asn Gln Ile Val Lys Glu Cys

1 5 10

Sequence Number : 45

Sequence Length : 10

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : N-terminal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Thr Phe Ala Tyr Lys Ile Asp Gly Asn Glu

1 5 10

Sequence Number : 46

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Leu Gly Pro Tyr Phe Ser Gln

1

5

Sequence Number : 47

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Asp Val Phe Val Tyr Asp Gly

1

5

Sequence Number : 48

Sequence Length : 19

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source

Organism : *Sulfolobus solfataricus*

Strain : KM1

Sequence

Tyr Asn Arg Ile Val Ile Ala Glu Ser Asp Leu Asn Asp Pro Arg Val

1

5

10

15

Val Asn Pro

Sequence Number : 49

Sequence Length : 5

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : *Sulfolobus acidocaldarius*

Strain : ATCC 33909

Sequence

Leu Asp Tyr Leu Lys

1

5

Sequence Number : 50

Sequence Length : 17

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Lys Arg Glu Ile Pro Asp Pro Ala Ser Arg Tyr Gln Pro Leu Gly Val

1

5

10

15

His

17

Sequence Number : 51

Sequence Length : 9

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Lys Asp Val Phe Val Tyr Asp Gly Lys

1

5

Sequence Number : 52

Sequence Length : 9

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

His Ile Leu Gln Glu Ile Ala Glu Lys

1

5

Sequence Number : 53

Sequence Length : 10

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Lys Leu Trp Ala Pro Tyr Val Asn Ser Val

1 5 10

Sequence Number : 54

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Met Phe Ser Phe Gly Gly Asn

1 5

Sequence Number : 55

Sequence Length : 14

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Asp Tyr Try Tyr Gln Asp Phe Gly Arg Ile Glu Asp Ile Glu

1 5 10

Sequence Number : 56

Sequence Length : 7

Type of Sequence : Amino acid

Strandedness : Single

Topology : Linear

Molecule Type : Peptide

Type of Fragment : Internal fragment

Original Source : Sulfolobus acidocaldarius

Strain : ATCC 33909

Sequence

Lys Ile Asp Ala Gln Trp Val

1 5

Sequence Number : 57

Sequence Length : 18

Type of Sequence : Nucleic acid

Molecule Type : Other nucleic acid (Synthesized DNA)

Sequence

AGCWAGKAGM TAYCARCC

Sequence Number : 58

Sequence Length : 24

Type of Sequence : Nucleic acid

Strandedness : Single

Topology : Linear

Molecule Type : Other nucleic acid (Synthesized DNA)

Sequence

YTTHCCATCR TAWACRAAWA CATC

[illegible]